A Test of Responsive Virtual Human Technology as an Interviewer Skills Training Tool

by
Michael W. Link, Ph.D., Polly P. Armsby, BA, Robert Hubal, Ph.D, and Curry I. Guinn, PhD.
RTI International

Abstract

Research on survey non-response suggests that advanced communication and listening skills are among the best strategies telephone interviewers can employ for obtaining survey participation, allowing them to identify and address respondents' concerns immediately with appropriate, tailored language. Yet, training on interaction skills is typically insufficient, relying on role-playing or passive learning through lecture and videos. What is required is repetitive, structured practice in a realistic work environment.

This research examines acceptance by trainees of an application based on responsive virtual human technology (RVHT) as a tool for teaching refusal avoidance skills to telephone interviewers. The application tested here allows interviewers to practice confronting common objections offered by reluctant sample members. Trainee acceptance of the training tool as a realistic simulation of "real life" interviewing situations is the first phase in evaluating the overall effectiveness of the RVHT approach. Data were gathered from two sources -- structured debrief questionnaires administered to users of the application, and observations of users by researchers and instructors. The application was tested with a group of approximately fifty telephone interviewers of varying skill and experience levels. The research presents findings from these acceptance evaluations and discusses users' experiences with and perceived effectiveness of the virtual training tool.

Contact Information:
Michael W. Link
RTI International
3040 Cornwallis Rd.
RTP, NC 27709
(919) 990-8462
Link@rti.org

Paper prepared for presentation at the 2002 Annual Conference of the American Association for Public Opinion Research, St. Pete, FL.

A Test of Responsive Virtual Human Technology as an Interviewer Skills Training Tool[1]

by
Michael W. Link, Ph.D., Polly P. Armsby, BA, Robert Hubal, Ph.D, and Curry I. Guinn, PhD.

Responsive Virtual Human Technology (RVHT) involves the use of natural language processing and an emotive behavioral engine to produce natural, interactive dialogues with intelligent, emotive virtual-reality (VR) agents. RVHT has great potential for use in training interaction skills, such as those required for effective survey interviewing. However, our understanding of how people interact with responsive virtual humans (a.k.a. intelligent agents) is quite limited. Better understanding requires employing RVHT in training applications and conducting systematic use, usability, perception, and training-effectiveness assessments. Important questions yet to be answered include:

- Do intelligent agents make learning more accessible?
- How willing are students to accept intelligent agents as interactive partners in learning?
- What skills can be acquired, practiced, and validated using RVHT?
- What is involved in providing a convincing simulation of human interaction, realistic enough for the student to suspend disbelief and acquire skills that will transfer to a "live" environment?

Users' interactions with RVHT applications are little studied and poorly understood. The research presented here (and the larger research program from which it is drawn) provides an initial assessment of some of the issues associated with user interface design, user acceptance of computer-based training, and perceptions of the effectiveness of the training tool. As part of this assessment, usability assessments were conducted using instructor observations, moderated focus groups, and a structured questionnaire. The assessment involved the use of an RVHT-based training tool for refusal avoidance at the outset of a telephone interview. Approximately fifty telephone interviewers of varying experience levels, ages, genders, races, and educational backgrounds took part in the assessment.

[1] This work was supported by a research grant from the National Science Foundation (Grant No. EIA-0121211) and by a Strategic Capability Development Award from RTI (R9898-002).

Background

Intelligent agents are being used in fields as diverse as computer generated (military) forces (Hill, et. al., 1998), manufacturing (Regian, Shebilske, and Monk, 1992), medicine (Miksch, Chang, and Hayes-Roth, 1996), and theater (Loyall and Bates, 1997; Lundeberg and Beskow, 1999). Intelligent agents have not been employed in training on interaction skills, although such skills are critical in a number of fields. Therefore, advanced technologies for training these "soft skills" can be a considerable asset in training. There remain, however, questions that must be answered if intelligent agents are to reach the level of sophistication required for robust interaction skills training.

Interaction skills training is certainly a new educational area in which to apply advances in information technology, such as virtual reality (VR) and agent technology. To date, VR has been shown to be effective for equipment training (Adams, 1996), maintenance training (Barnett, Helbing, Hancock, Heininger, and Perrin, 2000), simulation of military field exercises (Shlechter, Bessemer, and Kolosh, 1992), and maneuvers (Magee, 1995), and acquisition of spatial knowledge (Regian, Shebilske, and Monk, 1992). It can be used for interaction with unobservable processes or abstract concepts (Dede, Salzman, and Loftin, 1996), tasks that are costly or dangerous to perform (Loftin and Kenney, 1994), and for gaining situation awareness (Maggart and Hubal, 1998). VR systems have become steadily smaller, faster, cheaper, and easier to use (Psotka, 1995). RTI International has integrated a spoken natural language assistant with a VR-based maintenance training environment to enhance ease of use and facilitate instruction (Guinn and Montoya, 1998). Other relevant research efforts in enabling spoken interaction with virtual humans include work done at the University of Pennsylvania (Badler, Phillips, and Webber, 1993), MIT Media Lab (Cassell and Vilhjálmsson, 1999), University of Southern California (Lindheim and Swartout, 2001), and Oregon Graduate Institute (Cole et al, 1999; Massaro et. Al, 1998).

RVHT is a relatively recent advance in training technology. Few researchers have begun integrating emotion models with agents (Becheiraz and Thalmann, 1998; Elliott, 1993; Gratch, 2000; and Klein, 1998), and none for interaction training. Portraying emotions in a virtual human, it is argued, requires clearly defined emotional states, action that shows thought processes, and accentuation to reveal feelings (Bates, 1994). In general, lifelike "pedagogical agents" can lead to improvements in problem-solving ability and can engage and motivate trainees (Johnson, Rickel, and Lester, 2000; Lester et. al, 1997). Most importantly, RVHT can open entirely new capabilities for computer-based training of interpersonal skills, and can provide the benefits of reduced training costs, individualized tutoring, and greater student convenience that are associated with computer-based training (Field, et. al., 1999).

Today, interaction skills training usually relies on peer-to-peer role playing or passive learning through videos. These approaches lead to a critical training gap, because the students are limited in the practice time and the variety of scenarios that they encounter. Nevertheless, it is exactly this practice that leads to significant on-the-job benefits.

Table 1 (adapted from Hubal, et al. 2000) presents a comparison of approaches to interaction skills training. Constraints imposed by the current approach include insufficient time in the classroom to conduct effective practice sessions, forced and unrealistic role-playing exercises, and little time or ability for individual feedback and coaching to trainees from the instructor. By using virtual humans to simulate realistic interactions, RVHT increases the amount of time trainees spend acquiring and practicing critical skills, reduces passive learning (information and skills are retained better through active learning), improves the realism of practice sessions, and enables intelligent tutoring (Graesser et al, 2000).

We stress that using virtual humans as interaction partners has disadvantages as well as advantages. Most importantly, the current state-of-the-art does not produce fully realistic conversational partners. Advances in utilizing natural language dialog features and behavior models will add tremendously to the realism. From a larger perspective, though, one must understand that virtual training is simply one component of training. Just as a trainee must "skin his/her knuckles" on actual machines in validating maintenance and diagnostic skills, so a trainee must interact with people in validating interaction skills (Helms, Hubal, Triplett, 1997). Virtual environments, though, offer advantages in reliability, repetitiveness, flexibility, throughput, and distribution that lead directly to overall cost-effectiveness of training (Field, et al, 1999).

Table 1. Comparison of Training Approaches

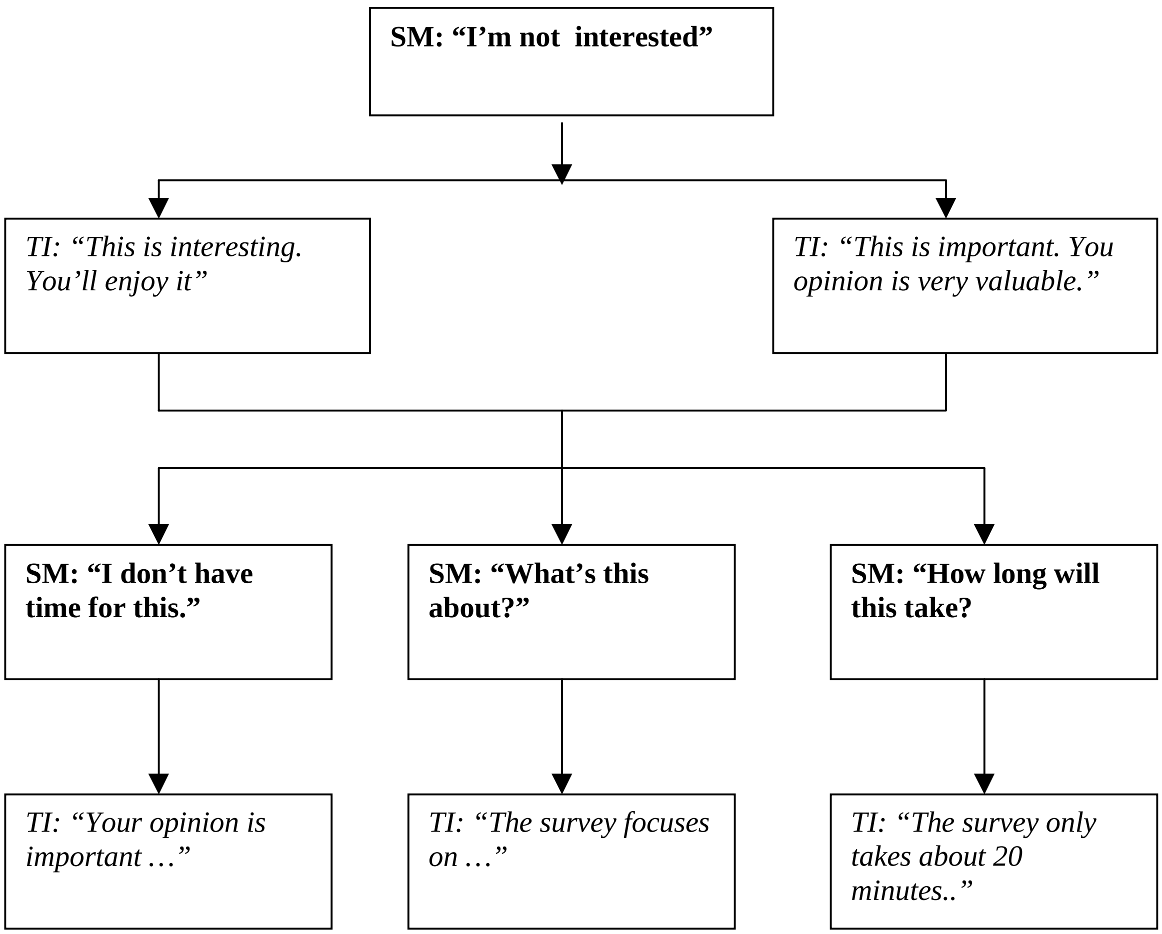
Role	Traditional Approach	Role-player	RVHT Approach	Role-player
Trainee (e.g., medical practitioner, police recruit, survey interviewer)	Student's ability to learn dependent on: ❑ relevance of role-play scripts, ❑ time available during training to conduct role-plays or mock interviews, ❑ acting ability of role-play Partner, ❑ observations made by role-play Partner and/or by Instructor.	Student	Student's ability to learn enhanced by: ❑ using numerous age-appropriate role-play or mock interview scripts, for more practice of critical skills, ❑ interacting with different virtual role-play partners, ❑ knowing that actions are observed and tracked, ❑ ability to replay interaction.	Student
Conversation Partner (e.g., patient, mentally disturbed consumer, household respondent)	❑ Partner must be present, available. ❑ Partner must act out a role that s/he will not always understand (non-essential learning activity). ❑ Partner is of a specific gender/age/ethnicity, limiting realism of practice.	Other person (e.g., actor, other student, Instructor)	❑ Ability to simulate conditions impossible with a human. ❑ Standardization of responses. ❑ Different virtual partners of gender/age/ethnicity and having different personalities.	Virtual human
Observer/ Evaluator	❑ Role-play Partner must take on second role, again a role not taken in live environment. ❑ Role-play Partner, if other student, is in passive learning mode.	Other person	❑ Ability to track all interactions with virtual role-play partner for use in feedback, guidance, assessment. ❑ Knowledge of all characteristics of virtual partners.	Second virtual human
Coach/Tutor	❑ Instructor must rely on role-play Partner for assessment of Student when not actually witnessing interaction. ❑ Only means of replaying interaction is through video, requiring an additional person and equipment.	Instructor or Supervisor	❑ Virtual tutor has ability to guide learning as it occurs. ❑ Instructor can use automatically collected interaction information for assessment & replay, as well as actually witness interaction. ❑ Instructor can convey "what-if" scenarios.	Second virtual human Instructor

Mechanics of the Training Application

One of the most difficult skills for a telephone interviewer to learn – and for an instructor to teach – is gaining cooperation from sample members and avoiding refusals. In telephone interviewing in particular, the first 30 seconds on the telephone with a sample member is crucial. Sample members almost automatically turn to phrases such as, “I don't do surveys,” “I don't have time,” “I'm just not interested” to avoid taking part in surveys. Non-response research suggests that the best approach to obtaining participation is for the interviewer to immediately reply with an appropriate, informative, tailored response (Camburn, Gunther-Mohr, & Lessler, 1999; Groves & Couper, 1998; Groves, 2002). How can the interviewer learn and then practice those responses before the survey begins, without creating more refusals during their first few weeks at work by being placed on the telephone unprepared?

The approach tested here involves the use of an RVHT-based application to simulate the environment an interview faces during the first 30 to 60 seconds of a telephone interview. The application allows interviewers to practice their skills in gaining cooperation in a self-paced, realistic environment. The software is designed such that interviewers begin with an introduction and then need to respond to a series of these objections or questions raised by the "virtual respondent." The interviewer's responses are captured electronically and processed by a natural language speech processor. Based on the content of the interviewer's speech, the software launches another objection/question or ends the conversation by either granting the interview or hanging-up the telephone (see Figure 1).

Figure 1. Example of Dialogue Flow



The application uses speech recognition and a behavior engine (for determining the intelligent agent's emotional state) to produce natural dialogues with the trainees. The speech recognizer uses a basic dictionary of common words as well as a specific dictionary for each turn of a conversation. The specific dictionary consists of up to 200 words based on behavioral observations of real world events. These specific dictionaries are dynamic, therefore, changing with each turn of the conversation. During the development of the application tested here, the researchers monitored live interviews and behavior coded the responses of interviewers and sample members. These behavioral observations were then modeled, using the dictionaries and the emotional state behavior engine. Thus the specific dictionaries created for capturing responses from an interviewer to a respondent who said, “I’m too busy” in a harsh tone varied somewhat from the dictionaries created for when the respondent gave the same objection but in a softer, more reasoned tone. As trainees used the application, the emotional state of the virtual respondent varied from scenario to scenario, thus giving trainees exposure to an array of objections and emotional states. The scripts launched by the RVHT program were recorded in both a male and a female voice to add variety to the program. In all a total of six basic objections were recorded in four different tones of voice for both a male and female virtual respondent. Thus a total of 48 different practice scenarios could be offered to the trainees.

Assessment of the RVHT-based Interviewer Training Application

A primary goal of the overall research program of which this study is a part is to determine if RVHT can be an effective technology for interaction training across a broad spectrum of ethnic and socioeconomic backgrounds, jobs, and job levels. In particular, we investigate whether users find RVHT interactions accessible and acceptable. The effectiveness of this technology depends upon its ability to provide appropriate learning experiences, its ability to engage the trainee, and its acceptability to disparate users.

An "accessible" user interface is one that is easy to learn and easy to use, and can result in measurable goals such as decreased learning time and greater user satisfaction (i.e., acceptance) (Weiss, 1993). Characteristics of easy to learn and easy to use interfaces have been described as having navigational and visual consistency, clear communication between the user and application, appropriate representations, few and non-catastrophic errors, task support and feedback, and user control (Nielsen, 1993; Norman, 1993; Shneiderman, 1992; Weiss, 1993).

The assessment provided here of the interviewer training module is based on researcher / instructor observations, and user debriefings in the form of a questionnaire. Empirical data were collected on users' observed ability to interact with the application as well as their perception of the interaction. The training application was tested with a group of approximately 50 telephone interviewers of varying ages, races, experience and education levels. Trainees who participated in the assessment used the application to practice communication and thinking skills required with real conversation partners. These skills involve the use of adaptive strategies, listening and responding to the other's concerns.

To evaluate the *accessibility* of the application we focused on the following:

- Do users understand the basic features of the application?
- Are users able to complete each task and exit the application?
- Do users understand where they are in the application?
- Are different users (e.g., based on ethnicity, job level, and education level) equally able to use the application?

Instructor/researcher observation was used to assess more directly the interaction between the user and the training application, addressing questions such as:

- When there are problems (e.g., the virtual human seems to respond inappropriately), what are user reactions?
- Are inappropriate responses due to a programming error, misunderstanding in the interaction, or incorrect user behavior?
- What knowledge engineering improvements will lead to better recovery by the application when inappropriate responses occur?

Analysis of these questions will provide clues as to how smoothly the application runs, or when and why difficulties arise in its use.

The question of whether and why participants "accept" or "reject" the virtual training environment is also central to this research. To evaluate *acceptance* of the application by the trainees, we debriefed participants using a structured questionnaire and moderator-facilitated focus groups to gauge reactions and engagement in the application. In particular we are interested in the following:

- Are the virtual humans realistic enough for the users? Why or why not?
- How fast and accurate is the speech recognition?
- When recognition is inaccurate, does the application respond reasonably?
- Overall, do the users "buy into" the virtual environment?
- Could trainees detect changes in the emotive states of the virtual human using only audio cues?
- Did the trainee perceive any gains in skills from using the application?
- Would they use the application again and/or recommend it's use by others?

While some of these acceptance measures may be particular to the specific application tested, most help in gaining a general understanding of user satisfaction and affect with RVHT.

As part of the evaluation process, data were collected using a questionnaire filled out by the interviewers and notes made by instructors and researchers who observed the training sessions. The questionnaire asked questions related to users' perceptions of the realism of the interactions with the "virtual human," ease of use of the software, the perceived effectiveness of the training sessions, and some basic background characteristics of the users. In all, a diverse group of 48 interviewers filled-out the questionnaires (96% of the software users). A breakdown of some of the demographic characteristics of this set of users is provided on Table 2.

Finally, each training session was observed by either the researchers or training instructors, who made notes of their observations. These observations are included as part of the analysis.

Findings

The questions posed to the interviewers were designed to assess their perceptions and experiences in using the RVHT training tool in four basic areas: ease of use of the software, realism of the training environment, impact on skill development, and desire to recommend or use the software again. Although this is the first detailed look at how users interact emotive intelligent agents for soft-skills development, we can formulate some hypotheses regarding how different types of users might respond based on how users generally differ in their use and acceptance of other computer-based tools. For example, we might expect to find that trainees who are younger, have more education, and are more comfortable using computers in general to have fewer difficulties in using the system. Likewise, we might expect that more experienced interviewers might not find the training tool as useful as inexperienced interviewers because the more experienced interviewers will have already developed and honed their refusal avoidance skills (a supposition that mirrors the finding of Groves, 2002). To examine possible differences in accessibility and acceptance of the program, we cross-tabulated all of the closed-ended questions in the questionnaire with the demographic variables listed on Table 2. Significant differences are noted in the text.[2]

Table 2. Demographics of RVHT Trainees

Characteristic	**N**	**%**
Sex		
Male	12	25%
Female	36	75%
Education		
High School/GED	2	4%
Some College	12	25%
Four Year Degree	25	52%
Advanced Degree	9	19%
Age		
18-21	7	15%
22-29	17	35%
30-39	8	17%
40-49	7	15%
50+	9	18%
Race		
African-American	34	70%
White	7	15%
Hispanic	7	15%
Experience		
< 1,000 hours	19	40%
1,000 – 1,999 hours	17	35%
2,000+ hours	12	25%
Comfort with Keyboard		
Slow-touch typing	15	31%
Fast-touch typing	33	69%

[2] Because of the small number of observations (N=48) we also created dichotomous variables for both the dependent variables (collapsing scales where possible) and independent variables (collapsing or combining variables with 3 or more values). These variables were also examined to determine if significant differences among subgroups could be identified.

Ease of Use of the Application

Training software should be accessible to users; that is, it should be relatively easy to use. As shown on Table 3, users of the RVHT software seemed to find it very accessible to use, with 84% indicating the software was either extremely easy or very easy to use (52% extremely, 31% very, 13% somewhat, 4% not too, 0% not at all). Nearly everyone found the written instructions (96%) and the verbal instructions (98%) that accompanied the training to be clear and accurate. Only eight (17%) of the 48 trainees indicated that they required additional assistance to use the training software (after the initial training received by all trainees).

The only significant difficulty encountered by the users were "insufficient memory" errors received on some of the training stations. The application does, at times, use up considerable CPU memory. Once the machines were adjusted to handle the software memory requirements, the error messages were no longer an issue.

Table 3. Interviewer's Evaluation of the RVHT Training Software

	Extremely	Very	Somewhat	Not Too	Not At All
In general, how easy was the application to use?	52.1% (25)	31.3% (15)	12.5% (6)	4.2% (2)	0 % (0)
In general, how realistic did you find the overall conversation with the "virtual respondent"?	2.1% (1)	14.6% (7)	43.8% (21)	16.7% (8)	22.9% (11)
In general, how realistic did you find the objections, concerns, questions posed by the "virtual respondent"?	12.5% (6)	35.4% (17)	39.6% (19)	8.3% (4)	4.2% (2)
How easily could you determine the "virtual respondent's" emotional state or attitude based on the tone of his/her voice?	22.9% (11)	43.8% (21)	29.2% (14)	4.2% (2)	0% (0)
How easily could you determine the "virtual respondent's" emotional state or attitude based on the words used or objectives raised by him/her?	8.3% (4)	54.2% (26)	27.1% (13)	10.4 % (5)	0% (0)

Realism of the Training Environment

The promise of RVHT-based training tools is that they can simulate a "real" environment, thereby allowing trainees repetitive practice in conditions that are as close as possible to what they will encounter on the job. For this particular application, the "virtual respondent" needed to mirror the behaviors and emotions of real respondents encountered when doing live interviewing. This means delivering an array of objections to the trainees in different tones of speech and emotional levels in a fast-paced manner. Interviewers were asked a series of questions to try to assess how well they accepted the virtual environment as a substitute for real work conditions. In other words, do they “buy-into” the virtual environment?

The answer is somewhat mixed. In general, trainees did not find the virtual environment to be realistic and they cited two primary reasons: the slowness of the response of the “virtual respondent” and the limited number of different objections/questions offered by the “virtual respondent.” They did, however, find the responses that were offered to be realistic and stated that they could detect and respond to changes in tone and emotional cues offered by the “virtual respondents.” A majority of the trainees also indicated that they felt the sessions helped them to improve their skills needed at the outset of an interview either somewhat or a lot.

When asked, *In general, how realistic did you find the overall conversation with the 'virtual respondent*,' 17% said they thought it was extremely or very realistic, 44% said it was somewhat realistic, 17% not too realistic and 23% not at all realistic (see Table 3). Slowness of the “virtual respondents” in replying (due to the lag caused by the speech recognizer as it interpreted the interviewer's responses and determined the next script to launch) was the primary problem cited by interviewers. Over three-quarters (77%) of the users felt the response time was too slow (4% felt it was too fast and 19% indicated the speed was just right). Perhaps not surprisingly, trainees who describe themselves as “fast-touch typists” were more likely than those who indicated they were “slow-touch typists” to say the response time was too slow (82% fast-touch vs 67% slow-touch; $p < .08$ chi-sq). Interviewers who are more comfortable at a keyboard and who, it can be surmised, tend to get through an interview faster were the ones most put-off by the perceived slowness of the response time.

The trainees were, however, more positive when evaluating the realism of the objections and questions offered by the “virtual respondent.” A plurality (48%) indicated that the content of what was said was either extremely or very realistic, with 40% saying it was somewhat realistic, 8% not too realistic, and 4% not at all realistic. They also felt it was relatively easy to determine the emotional state of the virtual respondent based on the tone of voice they heard (23% extremely easy, 44% very easy, 29% somewhat easy, and 4% not too easy; no one indicated that they could not determine the avatar's emotional state from the tone of the "virtual human's" voice). Likewise, the content of the speech used by the avatar was also a good cue to trainees as to the "virtual human's" emotional state: 8% extremely easy to tell, 54% very easy, 27% somewhat easy, 10% not too easy, 0% not at all easy.

Being able to recognize changes in the emotional state of the virtual respondent changed – at least in the minds of many trainees – how the interviewer approached the situation. Nearly 60% indicated that they behaved differently in the practice scenario based on the tone of the virtual respondent's voice. Interestingly, a higher percentage of women than men reported reacting differently to the changing tone of the avatar's voice (women 67% v. men 33%, $p < .04$ chi-sq.). Similarly, 54% said they treated the situation differently based on the actual words used by the avatar in expressing a concern or voicing an objection. There were, however, no differences between men and women on this question. When asked how they behaved differently, interviews said they tended to soften and take a more conciliatory tone when the virtual respondent seem to grow more hostile or angered, and they mirrored the tone when the virtual respondent seemed more pleasant. Likewise, they reported tailoring the content of their responses to try to meet the objections or questions of the virtual sample member rather than simply moving forward with their script. It seems, therefore, that the both the content of the objections raised by the virtual respondent and the emotional behavior of the "virtual human" were generally accepted by the trainees and caused them to react differently within the various training scenarios.

When asked in an open-ended format to list some of the problems with the realism of the software, many cited the slowness and others indicated that the limited number of objections raised by the virtual respondent made the sessions less realistic than what they encounter on the telephone. Because this was the first iteration of the software, a conscious decision was made at the design phase to maintain a limited set of six main objections and questions ("I'm not interested," "I'm too busy," "What is the survey about?", "I don't have time right now," "How was I selected?", and "How long will this take?"). These six responses, however, were recorded in four different tones of voice (ranging from calm to upset) and recorded in both a male and a female voice. A total of 48 possible practice scenarios was, therefore, actually possible (6 responses * 4 tones of voice * 2 sexes). It appears, however, that while the interviewers do recognize and react to the different emotional cues they obtain from the different scenarios, they don't necessarily process these as being very distinct. They focus more on the actual content of the argument (regardless of the tone of voice or whether the voice is a male or female) when considering how diverse the scenarios offered are. In designing future versions of the software this will need to be considered to increase interviewer acceptance of the training tool as a realistic simulation of the environment within which they must work.

Impact on Skill Development

The purpose for allowing trainees to operate within a virtual environment is to allow them to develop and hone essential skills before entering the "real" environment, thereby reducing the amount of "on the job" skill development required. For telephone interviewers, this means an opportunity to practice their skills at gaining cooperation at the outset of an interview. Practice in a virtual environment, it is hoped, will allow interviewers – particularly new interviewers – to develop, practice, and hone these skills before getting on the telephone. New interviewers can do considerable damage at the outset of a telephone study, generating a large number of refusals as they gain comfort and confidence on the telephone. If practice within a virtual environment at the beginning of a project can reduce the numbers of initial refusals even modestly, then the training program will have value. While longer-term assessments of the effectiveness of the RVHT software will need to include examination of more objective measures of improved performance, this preliminary assessment focused on the user's assessment of the impact of the training on their own skill development.

Trainees were asked to evaluate if they thought the RVHT software increased their abilities in six different areas (see Table 4). Nearly three-quarters of the trainees felt that the practice sessions increased a lot or somewhat their ability to respond to questions and concerns by sample members. Approximately 56% felt it helped them a lot or somewhat in better gaining respondent cooperation at the outset of an interview. Likewise, over half felt it helped in their ability to adapt to differences in respondents' tone or voice or perceived moods and to adapt to differences in the speed and pace of different sample members' speech. About half of the trainees also thought that the sessions helped them a lot or somewhat in avoiding refusals at the outset of an interview.

Table 4. Interviewer's Perceptions of Effectiveness of RVHT Training Software

	A Lot	Somewhat	A Little	Not at All
Respond to questions / concerns raised by sample members	25.0% (12)	47.9% (23)	16.7% (8)	10.4% (5)
Better gain respondent cooperation during the first seconds of a call	25.0% (12)	31.3% (15)	29.2% (14)	14.6% (7)
Enhance your ability to adapt to differences in respondents' tone/mood	25.0% (12)	29.2% (14)	29.2% (14)	16.7% (8)
Think on your feet	20.8% (10)	39.6% (19)	27.1% (13)	12.5% (6)
Enhance your ability to adapt to differences in respondents pace of speaking	18.8% (9)	33.3% (16)	27.1% (13)	20.8% (10)
Avoid refusals at the outset of an interview	16.7% (8)	35.4% (17)	31.3% (15)	16.7% (8)

Once again, while more objective measures of increased ability to gain cooperation from sample members are needed in the longer-term evaluation of this training tool, it does appear that trainees perceive an increase in their ability to deal with various facets of the opening of an interview as a result of their training sessions.

Would They Use The RVHT Training Tool Again?

An effective training tool is also one that trainees should enjoy using, would use again, and recommend to others (see Table 5). Approximately two-thirds (65%) of the users said that they found using the RVHT software to be fun and enjoyable. Interestingly men were significantly more likely than women to say that they found the sessions to be enjoyable (92% men vs. 56% women, $p < .05$ chi-sq). Nearly three-quarters (73%) said they would like to use the software again. In addition, 83% said they would recommend the program as a training tool for other interviewers. In open-ended responses, a number of interviewers indicated that it would be a very good practice vehicle for new or less experienced interviewers.

Table 5. Recommendation for Future Use of RVHT Training Tool

Assessment Questions:	**Yes**	**No**
Would you recommend the RVHT program as a training tool for other interviewers?	83% (40)	17% (8)
Would you like to use the RVHT program again as a training tool?	73% (35)	27% (13)
Was using RVHT fun and enjoyable?	65% (31)	35% (17)

Conclusions

This initial assessment of an RVHT-based training tool for telephone interviewers provides some valuable insights into how trainees access and accept virtual environments as practice labs and "virtual humans" as training partners. There were aspects of the training program that interviewers clearly liked, such as the ability to do repeated practice of frequently asked questions, being able to distinguish different emotional states from the tone of voice and speech content of the virtual respondent, and the opportunity to learn to think on their feet in a simulated environment before being placed into a live interviewing situation.

There were also aspects that the interviewers did not like, such as the slowness of the response of the virtual respondent and the perceived lack of variety in the scenarios that were presented. This provides constructive feedback for the engineering and improvement of the software. Adding additional scenarios is a relatively easy process, involving research into the "normal" flow of such scenarios and simple scripting and programming. The responsiveness issue is a more fundamental matter, reflecting the current state-of-the-art in speech recognition. For virtual training partners to be more readily accepted, the underlying speech recognition technology needs to be improved, providing quicker, more efficient processing of the input from interviewers and more rapid launching of responses by the virtual respondent. While our research focused on a specific training application, the results have implications for a broader range of training and educational RVHT-based tools. The lessons learned here can be used to inform the development of tools in these other areas.

We do not anticipate RVHT-based training will replace instructor-led training, but we expect that combinations of RVHT-based training and instructor-led training will significantly reduce training development costs (with new development tools) and training delivery costs, while increasing trainee throughput and maintaining training effectiveness and consistency. As an additional return-on-investment, RVHT-based training can provide inexpensive, focused sustainment (i.e., refresher) training. We feel it is important to continue to investigate more robust and effective RVHT models and more efficient means of creating the models, to better understand user preferences and acceptance of RVHT, and to determine how best to use RVHT in combination with other training methods to provide cost-effective training on critical interaction skills.

References

Adams, N. (1996). A Study of the Effectiveness of Using Virtual Reality to Orient Line Workers in a Manufacturing Environment. Motorola University, unpublished dissertation.

Badler, N.I., Phillips, C.B., & Webber, B.L. (1993). Simulating Humans: Computer Graphics, Animation, and Control. Oxford Univ. Press.

Barnett, B., Helbing, K., Hancock, G., Heininger, R., & Perrin, B. (2000). An Evaluation of the Training Effectiveness of Virtual Environments. Presented at the Interservice/Industry Training, Simulation and Education Conference. November 30, 2000, Orlando, FL.

Bates, J. (1994). The Role of Emotion in Believable Agents. Communications of the ACM, Special Issue on Agents, July, 1994.

Becheiraz, P., & Thalmann, D. (1998). A Behavioral Animation System for Autonomous Actors personified by Emotions, Proceedings of the First Workshop on Embodied Conversational Characters (WECC '98), Lake Tahoe, California.

Camburn, D.P., Gunther-Mohr, C., & Lessler, J.T., (1999). Developing New Models of Interviewer Training. International Conference on Survey Nonresponse, Portland, OR, October 28-31, 1999.

Cassell, J., & Vilhjálmsson, H.H. (1999). Fully Embodied Conversational Avatars: Making Communicative Behaviors Autonomous. Autonomous Agents and Multi-Agent Systems: 2, 45-64.

Cole, R., et. al. (1999). New tools for interactive speech and language training: Using animated conversational agents in the classrooms of profoundly deaf children. In Proceedings of ESCA-MATISSE ESCA/SOCRATES Workshop on Method and Tool Innovations for Speech Science Education, London, UK.

Dede, C., Salzman, M., & Loftin, B. (1996) ScienceSpace: Research on Using Virtual Reality to Enhance Science Education. In P. Carlson & F. Makedon (Eds.), Proceedings of the 1996 ED-MEDIA Conference (pp. 172-177). Charlottesville, VA: Association for the Advancement of Computers in Education.

Elliott, C. (1993). Using the Affective Reasoner to Support Social Simulations. In Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence, pp. 194-200, Chambery, France, August 1993. Morgan Kaufmann.

Field, S.S., Frank, G.A., Helms, R.F., & Hubal, R.C. (1999). Army Learning & Training Effectiveness Symposium. Final Report, March 30, 1999. Submitted to Battelle RTP Office, Subcontract # DAAH04-96-C-0086, Agreement #99020, Delivery Order #0366, Dated February 12, 1999.

Graesser, A., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & the Tutoring Research Group (2000). AutoTutor: A simulation of a human tutor. Journal of Cognitive Systems Research, 1,35-51.

Gratch, J. (2000). Modeling the Interplay Between Emotion and Decision-Making. Proceedings of the Ninth Conference on Computer Generated Forces & Behavioral Representation, May 16-18, 2000, Orlando, FL.

Groves, R. (2002). Principles and Practices in Non-response Reduction. Presentation at the 2002 Respondent Cooperation Workshop sponsored by the Council for Marketing and Opinion Research, New York, NY.

Groves, R., & Couper, M. (1998). Nonresponse in Household Interview Surveys. New York: John Wiley & Sons, Inc.

Guinn, C.I., & Montoya, R.J. (1998). Natural Language Processing in Virtual Reality, Modern Simulation and Training, pp. 44-55.

Helms, R.F., Hubal, R.C., & Triplett, S.E. (1997). Evaluation of the Conduct of Individual Maintenance Training in Live, Virtual, and Constructive (LVC) Training Environments and their Effectiveness in a Single Program of Instruction. Final Report, September 30, 1997. Submitted to Battelle RTP Office, Subcontract # TCN 97031, Delivery Order #0027, Dated April 16, 1997.

Hill, R., Chen, J., Gratch, J., Rosenbloom, P., & Tambe, M. (1998). Soar-RWA: Planning, Teamwork, and Intelligent Behavior for Synthetic Rotary Wing Aircraft, Proceedings of the Seventh Conference on Computer Generated Forces & Behavioral Representation, May 12-14, 2000, Orlando, FL.

Hubal, R.C., Kizakevich, P.N., Guinn, C.I., Merino, K.D., & West, S.L. (2000). The Virtual Standardized Patient-Simulated Patient-Practitioner Dialogue for Patient Interview Training. In J.D. Westwood, H.M. Hoffman, G.T. Mogel, R.A. Robb, & D. Stredney (Eds.), Envisioning Healing: Interactive Technology and the Patient-Practitioner Dialogue, 133-138. IOS Press: Amsterdam.

Johnson, W.L., Rickel, J.W., & Lester, J.C. (2000). Animated pedagogical agents: face-to-face interaction in interactive learning environments. International Journal of Artificial Intelligence in Education, 11.

Klein, G. (1998). Sources of Power. MIT Press.

Lester, J., Converse, S., Kahler, S., Barlow, S., Stone, B., & Bhogal, R. (1997). The Persona Effect: Affective Impact of Animated Pedagogical Agents. In S. Pemberton (Ed.), Human Factors in Computing Systems: CHI'97 Conference Proceedings. 359-366. New York: ACM Press.

Lindheim, R., & Swartout, W. (2001). Forging a New Simulation Technology at the ICT. Computer, 34(1), 72-79.

Loftin, R.B., & Kenney, P.J. (1994). Virtual Environments in Training: NASA's Hubble Space Telescope Mission. 16th Interservice/Industry Training Systems & Education Conference, Orlando, FL.

Loyall, A.B., & Bates, J. (1997). Personality Rich Believable Agents That Use Language. Proceedings of the First International Conference on Autonomous Agents, February 1997, Marina del Rey, CA.

Lundeberg, M., & Beskow, J. (1999). Developing a 3D-Agent for the August Dialogue System. Proceedings from AVSP '99, Santa Cruz, CA.

Magee, L.E. (1995). Virtual Reality and Distributed Interactive Simulation for Training Ship Formation Manoeuvres. Proceedings of the 36th NATO Defense Research Group (DRG) Seminar.

Maggart, L.E., & Hubal, R.C. (1998). A Situation Awareness Model. In S.E. Graham & M.D. Matthews (Eds.), Infantry Situation Awareness: Papers from the 1998 Infantry Situation Awareness Workshop. U.S. Army Research Institute: Alexandria, VA.

Massaro, D.W., Cohen, M.M., Beskow, J., Daniel, S., & Cole, R.A. (1998). Developing and Evaluating Conversational Agents. In Proceedings of Workshop on Embodied Conversation Characters (WECC), Lake Tahoe.

Miksch, S., Chang, K., & Hayes-Roth, B. (1996). An Intelligent Assistant for Patient Health Care. Knowledge Systems Laboratory Report No. 96-19, Stanford University.

Mills, K.C., Parkman, K.M., Smith, G.A., & Rosendahl, F. (1999). Prediction of driving performance through computerized testing: High-risk driver assessment and training. Transportation Research Record, 1689, 18-24.

Nielsen, J. (1993). Usability Engineering. Boston: Academic Press.

Norman, D.A. (1993). Things That Make Us Smart . Reading, MA: Addison-Wesley.

Psotka, J. (1995). Immersive Training Systems: Virtual Reality and Education and Training. Instructional Science 23, 405-31.

Regian, J.W., Shebilske, W.L., & Monk, J.M. (1992). Virtual reality: An instructional medium for visual-spatial tasks. Journal of Communication 42, 136-149.

Rousseau, D., & Hayes-Roth, B. (1997). Improvisational Synthetic Actors with Flexible Personalities. KSL Report #97-10, Stanford University.

Shlechter, T.M., Bessemer, D.W., & Kolosh, K.P. (1992). Computer-Based Simulation Systems and Role-playing: An Effective Combination for Fostering Conditional Knowledge. Journal of Computer Based Instruction 19, 110-114.

Shneiderman, B. (1992). Designing the User Interface: Strategies for Effective Human-Computer Interaction. Reading, MA: Addison-Wesley.

Weiss, E. (1993). Making Computers People-Literate. San Francisco: Jossey-Bass.